

V. PORTNER

1645

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/284,233A
DATE: 01/08/2001
TIME: 14:59:51

Input Set : A:\05649008.app
Output Set : N:\CRF3\01082001\I284233A.raw

P#13

ENTERED

RECEIVED

JAN 17 2001

TECH CENTER 1600/2900

5 <110> APPLICANT: Meyer, Thomas
7 Haas, Rainer
9 Zhengxin, Yan
11 Gomez-Duarte, Oscar
13 Lucas, Bernadette
17 <120> TITLE OF INVENTION: Helicobacter pylori live vaccine
21 <130> FILE REFERENCE: 100564-09008
25 <140> CURRENT APPLICATION NUMBER: 09/284,233A
27 <141> CURRENT FILING DATE: 1999-07-28
31 <150> PRIOR APPLICATION NUMBER: PCT/EP97/04744
33 <151> PRIOR FILING DATE: 1997-09-01
37 <150> PRIOR APPLICATION NUMBER: EP/96 116 337.5
39 <151> PRIOR FILING DATE: 1996-10-11
43 <160> NUMBER OF SEQ ID NOS: 9
47 <170> SOFTWARE: PatentIn version 3.0
51 <210> SEQ ID NO: 1
53 <211> LENGTH: 1557
55 <212> TYPE: DNA
57 <213> ORGANISM: Helicobacter pylori
61 <220> FEATURE:
63 <221> NAME/KEY: CDS
65 <222> LOCATION: (1)..(1554)
69 <400> SEQUENCE: 1
70 atg aca caa tct caa aaa gta aga ttc tta gcc cct tta agc cta gcg 48
71 Met Thr Gln Ser Gln Lys Val Arg Phe Leu Ala Pro Leu Ser Leu Ala
72 1 5 10 15
74 tta agc ttg agc ttc aat cca gtg ggc gct gaa gaa gat ggg ggc ttt 96
75 Leu Ser Leu Ser Phe Asn Pro Val Gly Ala Glu Glu Asp Gly Gly Phe
76 20 25 30
78 atg acc ttt ggg tat gaa tta ggt cag gtg gtc caa caa gtg aaa aac 144
79 Met Thr Phe Gly Tyr Glu Leu Gly Gln Val Val Gln Gln Val Lys Asn
80 35 40 45
82 ccg ggt aaa atc aaa gcc gaa gaa tta gcc ggc ttg tta aac tct acc 192
83 Pro Gly Lys Ile Lys Ala Glu Glu Leu Ala Gly Leu Leu Asn Ser Thr
84 50 55 60
86 aca aca aac aac acc aat atc aat att gca ggc aca gga ggc aat gtc 240
87 Thr Thr Asn Asn Thr Asn Ile Asn Ile Ala Gly Thr Gly Gly Asn Val
88 65 70 75 80
90 gcc ggg act ttg ggc aac ctt ttt atg aac caa tta ggc aat ttg att 288
91 Ala Gly Thr Leu Gly Asn Leu Phe Met Asn Gln Leu Gly Asn Leu Ile
92 85 90 95
94 gat ttg tat ccc act ttg aac act agt aat atc aca caa tgt ggc act 336
95 Asp Leu Tyr Pro Thr Leu Asn Thr Ser Asn Ile Thr Gln Cys Gly Thr
96 100 105 110
98 act aat agt ggt agt agt agt agt ggt ggt ggc gcc aca gcc gct 384
99 Thr Asn Ser Gly Ser Ser Ser Ser Gly Gly Gly Ala Ala Thr Ala Ala
100 115 120 125

BEST AVAILABLE COPY

RAW SEQUENCE LISTING

DATE: 01/08/2001

PATENT APPLICATION: US/09/284,233A

TIME: 14:59:51

Input Set : A:\05649008.app

Output Set: N:\CRF3\01082001\I284233A.raw

102	gct act act agc aat aag cct tgt ttc caa ggt aac ctg gat ctt tat	432
103	Ala Thr Thr Ser Asn Lys Pro Cys Phe Gln Gly Asn Leu Asp Leu Tyr	
104	130 135 140	
106	aga aaa atg gtt gac tct atc aaa act ttg agt caa aac atc agc aag	480
107	Arg Lys Met Val Asp Ser Ile Lys Thr Leu Ser Gln Asn Ile Ser Lys	
108	145 150 155 160	
110	aat atc ttt caa ggc aac aac aac acc acg agc caa aat ctc tcc aac	528
111	Asn Ile Phe Gln Gly Asn Asn Asn Thr Thr Ser Gln Asn Leu Ser Asn	
112	165 170 175	
114	cag ctc agt gag ctt aac acc gct agc gtt tat ttg act tac atg aac	576
115	Gln Leu Ser Glu Leu Asn Thr Ala Ser Val Tyr Leu Thr Tyr Met Asn	
116	180 185 190	
118	tcg ttc tta aac gcc aat aac caa gcg ggt ggg att ttt caa aac aac	624
119	Ser Phe Leu Asn Ala Asn Asn Gln Ala Gly Gly Ile Phe Gln Asn Asn	
120	195 200 205	
122	act aat caa gct tat gga aat ggg gtt acc gct caa caa atc gct tat	672
123	Thr Asn Gln Ala Tyr Gly Asn Gly Val Thr Ala Gln Gln Ile Ala Tyr	
124	210 215 220	
126	atc cta aag caa gct tca atc act atg ggg cca agc ggt gat agc ggt	720
127	Ile Leu Lys Gln Ala Ser Ile Thr Met Gly Pro Ser Gly Asp Ser Gly	
128	225 230 235 240	
130	gct gcc gca gcg ttt ttg gat gcc gct tta gcg caa cat gtt ttc aac	768
131	Ala Ala Ala Ala Phe Leu Asp Ala Ala Leu Ala Gln His Val Phe Asn	
132	245 250 255	
134	tcc gct aac gcc gcg aac gat ttg agc gct aag gaa ttc act agc ttg	816
135	Ser Ala Asn Ala Gly Asn Asp Leu Ser Ala Lys Glu Phe Thr Ser Leu	
136	260 265 270	
138	gtg caa aat atc gtc aat aat tct caa aac gct tta acg cta gcc aac	864
139	Val Gln Asn Ile Val Asn Asn Ser Gln Asn Ala Leu Thr Leu Ala Asn	
140	275 280 285	
142	aac gct aac atc agc aat tca aca ggc tat caa gtg agc tat ggc ggg	912
143	Asn Ala Asn Ile Ser Asn Ser Thr Gly Tyr Gln Val Ser Tyr Gly Gly	
144	290 295 300	
146	aat att gat caa gcg cga tct acc caa cta tta aac aac acc aca aac	960
147	Asn Ile Asp Gln Ala Arg Ser Thr Gln Leu Leu Asn Asn Thr Thr Asn	
148	305 310 315 320	
150	act ttg gct aaa gtt agc gct ttg aat aac gag ctt aaa gct aac cca	1008
151	Thr Leu Ala Lys Val Ser Ala Leu Asn Asn Glu Leu Lys Ala Asn Pro	
152	325 330 335	
154	tggt ctt ggg aat ttt gcc gcc ggt aac agc tct caa gtg aat gcg ttt	1056
155	Trp Leu Gly Asn Phe Ala Ala Gly Asn Ser Ser Gln Val Asn Ala Phe	
156	340 345 350	
158	aac ggg ttt atc act aaa atc ggt tac aag caa ttc ttt ggg gaa aac	1104
159	Asn Gly Phe Ile Thr Lys Ile Gly Tyr Lys Gln Phe Phe Gly Glu Asn	
160	355 360 365	
162	aag aat gtg ggc tta cgc tac tac ggc ttc ttc agc tat aac ggc gcg	1152
163	Lys Asn Val Gly Leu Arg Tyr Tyr Gly Phe Phe Ser Tyr Asn Gly Ala	
164	370 375 380	
166	ggc gtg ggt aat ggc cct act tac aat caa gtc aat ttg ctc act tat	1200

RAW SEQUENCE LISTING

DATE: 01/08/2001

PATENT APPLICATION: US/09/284,233A

TIME: 14:59:51

Input Set : A:\05649008.app

Output Set: N:\CRF3\01082001\I284233A.raw

```

167 Gly Val Gly Asn Gly Pro Thr Tyr Asn Gln Val Asn Leu Leu Thr Tyr
168 385 390 395 400
170 ggg gtg ggg act gat gtg ctt tac aat gtg ttt agc cgc tct ttt ggt 1248
171 Gly Val Gly Thr Asp Val Leu Tyr Asn Val Phe Ser Arg Ser Phe Gly
172 405 410 415
-----174 agt agg agt ctt aat gcg ggc ttc ttt ggg ggg atc caa ctc gca ggg-----1296
175 Ser Arg Ser Leu Asn Ala Gly Phe Phe Gly Gly Ile Gln Leu Ala Gly
176 420 425 430
178 gat act tac atc agc acg cta aga aac agc tct cag ctt gcg agc aga 1344
179 Asp Thr Tyr Ile Ser Thr Leu Arg Asn Ser Ser Gln Leu Ala Ser Arg
180 435 440 445
182 cct aca gcg acg aaa ttc caa ttc ttg ttt gat gtg ggc tta cgc atg 1392
183 Pro Thr Ala Thr Lys Phe Gln Phe Leu Phe Asp Val Gly Leu Arg Met
184 450 455 460
186 aac ttt ggt atc ttg aaa aaa gac ttg aaa agc cat aac cag cat tct 1440
187 Asn Phe Gly Ile Leu Lys Lys Asp Leu Lys Ser His Asn Gln His Ser
188 465 470 475 480
190 ata gaa atc ggt gtg caa atc cct acg att tac aac act tac tat aaa 1488
191 Ile Glu Ile Gly Val Gln Ile Pro Thr Ile Tyr Asn Thr Tyr Tyr Lys
192 485 490 495
194 gct ggc ggt gct gaa gtg aaa tac ttc cgc cct tat agc gtg tat tgg 1536
195 Ala Gly Gly Ala Glu Val Lys Tyr Phe Arg Pro Tyr Ser Val Tyr Trp
196 500 505 510
198 gtc tat ggc tac gcc ttc taa 1557
199 Val Tyr Gly Tyr Ala Phe
200 515
203 <210> SEQ ID NO: 2
205 <211> LENGTH: 518
207 <212> TYPE: PRT
209 <213> ORGANISM: Helicobacter pylori
213 <400> SEQUENCE: 2
215 Met Thr Gln Ser Gln Lys Val Arg Phe Leu Ala Pro Leu Ser Leu Ala
216 .1 5 10 15
219 Leu Ser Leu Ser Phe Asn Pro Val Gly Ala Glu Glu Asp Gly Gly Phe
220 20 25 30
223 Met Thr Phe Gly Tyr Glu Leu Gly Gln Val Val Gln Gln Val Lys Asn
224 35 40 45
227 Pro Gly Lys Ile Lys Ala Glu Glu Leu Ala Gly Leu Leu Asn Ser Thr
228 50 55 60
231 Thr Thr Asn Asn Thr Asn Ile Asn Ile Ala Gly Thr Gly Gly Asn Val
232 65 70 75 80
235 Ala Gly Thr Leu Gly Asn Leu Phe Met Asn Gln Leu Gly Asn Leu Ile
236 85 90 95
239 Asp Leu Tyr Pro Thr Leu Asn Thr Ser Asn Ile Thr Gln Cys Gly Thr
240 100 105 110
243 Thr Asn Ser Gly Ser Ser Ser Gly Gly Gly Ala Ala Thr Ala Ala
244 115 120 125
247 Ala Thr Thr Ser Asn Lys Pro Cys Phe Gln Gly Asn Leu Asp Leu Tyr
248 130 135 140

```

RAW SEQUENCE LISTING DATE: 01/08/2001
 PATENT APPLICATION: US/09/284,233A TIME: 14:59:51

Input Set : A:\05649008.app
 Output Set : N:\CRF3\01082001\I284233A.raw

```

251 Arg Lys Met Val Asp Ser Ile Lys Thr Leu Ser Gln Asn Ile Ser Lys
252 145 150 155 160
255 Asn Ile Phe Gln Gly Asn Asn Asn Thr Thr Ser Gln Asn Leu Ser Asn
256 165 170 175
259 Gln Leu Ser Glu Leu Asn Thr Ala Ser Val Tyr Leu Thr Tyr Met Asn
260 180 185 190
263 Ser Phe Leu Asn Ala Asn Asn Gln Ala Gly Gly Ile Phe Gln Asn Asn
264 195 200 205
267 Thr Asn Gln Ala Tyr Gly Asn Gly Val Thr Ala Gln Gln Ile Ala Tyr
268 210 215 220
271 Ile Leu Lys Gln Ala Ser Ile Thr Met Gly Pro Ser Gly Asp Ser Gly
272 225 230 235 240
275 Ala Ala Ala Ala Phe Leu Asp Ala Ala Leu Ala Gln His Val Phe Asn
276 245 250 255
279 Ser Ala Asn Ala Gly Asn Asp Leu Ser Ala Lys Glu Phe Thr Ser Leu
280 260 265 270
283 Val Gln Asn Ile Val Asn Asn Ser Gln Asn Ala Leu Thr Leu Ala Asn
284 275 280 285
287 Asn Ala Asn Ile Ser Asn Ser Thr Gly Tyr Gln Val Ser Tyr Gly Gly
288 290 295 300
291 Asn Ile Asp Gln Ala Arg Ser Thr Gln Leu Leu Asn Asn Thr Thr Asn
292 305 310 315 320
295 Thr Leu Ala Lys Val Ser Ala Leu Asn Asn Glu Leu Lys Ala Asn Pro
296 325 330 335
299 Trp Leu Gly Asn Phe Ala Ala Gly Asn Ser Ser Gln Val Asn Ala Phe
300 340 345 350
303 Asn Gly Phe Ile Thr Lys Ile Gly Tyr Lys Gln Phe Phe Gly Glu Asn
304 355 360 365
307 Lys Asn Val Gly Leu Arg Tyr Tyr Gly Phe Phe Ser Tyr Asn Gly Ala
308 370 375 380
311 Gly Val Gly Asn Gly Pro Thr Tyr Asn Gln Val Asn Leu Leu Thr Tyr
312 385 390 395 400
315 Gly Val Gly Thr Asp Val Leu Tyr Asn Val Phe Ser Arg Ser Phe Gly
316 405 410 415
319 Ser Arg Ser Leu Asn Ala Gly Phe Phe Gly Gly Ile Gln Leu Ala Gly
320 420 425 430
323 Asp Thr Tyr Ile Ser Thr Leu Arg Asn Ser Ser Gln Leu Ala Ser Arg
324 435 440 445
327 Pro Thr Ala Thr Lys Phe Gln Phe Leu Phe Asp Val Gly Leu Arg Met
328 450 455 460
331 Asn Phe Gly Ile Leu Lys Lys Asp Leu Lys Ser His Asn Gln His Ser
332 465 470 475 480
335 Ile Glu Ile Gly Val Gln Ile Pro Thr Ile Tyr Asn Thr Tyr Tyr Lys
336 485 490 495
339 Ala Gly Gly Ala Glu Val Lys Tyr Phe Arg Pro Tyr Ser Val Tyr Trp
340 500 505 510
343 Val Tyr Gly Tyr Ala Phe
344 515
347 <210> SEQ ID NO: 3

```

RECEIVED

JAN 17 2001

TECH CENTER 1800/2000

RAW SEQUENCE LISTING

DATE: 01/08/2001

PATENT APPLICATION: US/09/284,233A

TIME: 14:59:51

Input Set : A:\05649008.app

Output Set: N:\CRF3\01082001\I284233A.raw

```

349 <211> LENGTH: 1557
351 <212> TYPE: DNA
353 <213> ORGANISM: Helicobacter pylori
357 <220> FEATURE:
359 <221> NAME/KEY: CDS
361 <222> LOCATION: (1)..(1554)
365 <400> SEQUENCE: 3
366 atq ata aaa aag aat aga acg ctg ttt ctt agt cta gcc ctt tgc gct 48
367 Met Ile Lys Lys Asn Arg Thr Leu Phe Leu Ser Leu Ala Leu Cys Ala
368 1 5 10 15
370 agc ata agt tat gcc gaa gat gat gga ggg ttt ttc acc gtc ggt tat 96
371 Ser Ile Ser Tyr Ala Glu Asp Asp Gly Gly Phe Phe Thr Val Gly Tyr
372 20 25 30
374 cag ctg ggg caa gtc atg caa gat gtc caa aac cca ggc ggc gct aaa 144
375 Gln Leu Gly Gln Val Met Gln Asp Val Gln Asn Pro Gly Gly Ala Lys
376 35 40 45
378 agc gac gaa ctg gcc aga gag ctt aac gct gat gta acg aac aac att 192
379 Ser Asp Glu Leu Ala Arg Glu Leu Asn Ala Asp Val Thr Asn Asn Ile
380 50 55 60
382 tta aac aac aac acc gga ggc aac atc gca ggg gcg ttg agt aac gct 240
383 Leu Asn Asn Asn Thr Gly Gly Asn Ile Ala Gly Ala Leu Ser Asn Ala
384 65 70 75 80
386 ttc tcc caa tac ctt tat tcg ctt tta ggg gct tac ccc aca aaa ctg 288
387 Phe Ser Gln Tyr Leu Tyr Ser Leu Leu Gly Ala Tyr Pro Thr Lys Leu
388 85 90 95
390 aat ggt agc gat gtg tct gcg aac gct ctt tta agt ggt gcg gta ggc 336
391 Asn Gly Ser Asp Val Ser Ala Asn Ala Leu Leu Ser Gly Ala Val Gly
392 100 105 110
394 tct ggg act tgt gcg gct gca ggg acg gct ggt ggc act tct ctt aac 384
395 Ser Gly Thr Cys Ala Ala Ala Gly Thr Ala Gly Gly Thr Ser Leu Asn
396 115 120 125
398 act caa agc act tgc acc gtt gcg ggc tat tac tgg ctg cct agc ttg 432
399 Thr Gln Ser Thr Cys Thr Val Ala Gly Tyr Tyr Trp Leu Pro Ser Leu
400 130 135 140
402 act gac agg att tta agc acg atc ggc agc cag act aac tac ggc acg 480
403 Thr Asp Arg Ile Leu Ser Thr Ile Gly Ser Gln Thr Asn Tyr Gly Thr
404 145 150 155 160
406 aac acc aat ttc ccc aac atg caa caa cag ctg acc tac ttg aat gcg 528
407 Asn Thr Asn Phe Pro Asn Met Gln Gln Gln Leu Thr Tyr Leu Asn Ala
408 165 170 175
410 ggg aat gtg ttt ttt aat gcg atg aat aag gct tta gag aat aag aat 576
411 Gly Asn Val Phe Phe Asn Ala Met Asn Lys Ala Leu Glu Asn Lys Asn
412 180 185 190
414 gga act agt agt gct agt gga act agt ggt gcg act ggt tca gat ggt 624
415 Gly Thr Ser Ser Ala Ser Gly Thr Ser Gly Ala Thr Gly Ser Asp Gly
416 195 200 205
418 caa act tac tcc aca caa gct atc caa tac ctt caa ggc caa caa aat 672
419 Gln Thr Tyr Ser Thr Gln Ala Ile Gln Tyr Leu Gln Gly Gln Gln Asn
420 210 215 220

```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/284,233A

DATE: 01/08/2001
TIME: 14:59:52

Input Set : A:\05649008.app
Output Set: N:\CRF3\01082001\I284233A.raw